Journal of Global Research in Electronics and Communication
Volume 1, No. 1, January 2025
Available Online at: www.jgrec.info
RESEARCH PAPER

Improving Fault Detection Accuracy in Semiconductor Manufacturing with Machine Learning Approaches

Vikas Prajapati
Individual Researcher
Emai: Prajapati.vikas2707@gmail.com

***Abstract*—The production of semiconductors is among the most technologically complex and intricate industrial processes. Given the hundreds of steps involved in semiconductor production, a defective wafer detection system that enables early wafer identification is necessary. Virtual metrology (VM) and statistical process control (SPC) have been used to identify defective wafers. The purpose of this research is to examine how machine learning (ML) methods may be used to enhance the accuracy of semiconductor manufacturing defect detection. By leveraging the WM811K dataset, which includes over 800,000 wafer images with multiple defect categories, the research applies a Convolutional Neural Network (CNN) integrated with data augmentation to enhance model performance. The proposed CNN-AUG model effectively addresses challenges such as data imbalance and overfitting, yielding an accuracy 98.56%, precision 98.77%, recall 98.78%, and an F1-score 98.77%. Comparative analysis with VGG19 and XGBoost demonstrates the superior performance of CNN-AUG in capturing intricate spatial features and improving fault detection efficiency. The results highlight the potential of ML-based approaches for optimizing semiconductor manufacturing processes, reducing defects, and enhancing yield.**

Keywords—Semiconductor Manufacturing, Fault Detection, Machine Learning, WM811K Dataset, Data Augmentation.

I. INTRODUCTION

The detection of faults within semiconductor production stands essential for reaching optimal operational efficiency alongside maintaining superior product quality standards[1]. The detection of defects becomes challenging because semiconductor fabrication includes various complex multi-step processes that involve silicon wafer treatment through deposition and etching photolithography and ion implantation[2]. The use of highly automated precise equipment in clean environments fails to eliminate inevitable errors and defects which appear in wafer dies. The combination of Machine learning (ML) methodologies with sophisticated preprocessing techniques shows promise for improved manufacturing results and enhanced defect detection[3][4].

The semiconductor production technique oversees the complete series of manufacturing sequences for batches of 25 silicon wafers until integrated circuits (ICs) are functional[5]. Multiple process operations extending for months result in hundreds of disruptions including equipment deterioration coupled with maintenance routines and changes in environmental conditions[6]. The multiple stages of process variability appear at intra-wafer and inter-wafer and intra-batch and inter-batch levels due to such disruptions[7]. Large amounts of process equipment sensor data, including temperature and pressure, together with gas flow and power data, can be analyzed through ML models to solve current manufacturing difficulties[8]. As the data-driven technique provides live process supervision and promotes timely fault identification it attains the purpose of lowering operational variations and enhancing yield performance[9].

Manufacturers who merge ML into their semiconductor operations receive operational advantages in addition to enhanced fault identification[10]. The implementation of these system elements produces lower maintenance expenses, shorter equipment outages, longer spare part lifespan better operational safety and better total production output[11]. Industrial facilities can discover and address system problems in advance through predictive maintenance by detecting faults this enables them to avoid high maintenance expenses and maintain consistent output quality[12].

A. Motivation and Contributions of the Study

In semiconductor manufacturing, high yield, cost reduction, and product dependability all depend on accurate defect detection. Traditional fault detection methods often struggle with complex manufacturing processes, high-dimensional data, and subtle defect patterns. A viable alternative is ML, which uses sophisticated algorithms to sift through massive datasets in search of outliers and ways to enhance predicted accuracy. By integrating these techniques, manufacturers can enhance process efficiency, minimize defects, and achieve higher levels of automation. The ultimate goal of this research is to help make semiconductor manufacturing more dependable by investigating how well machine learning models work for problem detection. This research's principal contributions encompass:

- Employs the WM811K Dataset to examine wafer defect patterns, guaranteeing rigorous model training and assessment.
- Employs preprocessing methods include the management of absent data, the use of labelled datasets, and the application of data augmentation to enhance data quality and model generalization.
- Assesses various machine learning models, including CNN-AUG, VGG19, and XGBoost, to extract spatial

© JGREC 2025, All Rights Reserved

features, hierarchical representations, and advantages of ensemble learning for precise defect identification.
- Examines models using F1-score, Accuracy, Precision, and Recall to guarantee a comprehensive evaluation of detection resilience and effectiveness.

B. *Organization of the paper*

This paper explores improving fault detection in semiconductor manufacturing using machine learning. Section II reviews existing techniques, Section III examines machine learning's role, and Section IV covers data and preprocessing. Section V details the proposed CNN model, Section VI compares it with baseline models, and Section VII discusses results and challenges. Important results and directions for the future are presented in Section VIII.

II. Literature Review

The literature review on fault detection in semiconductor manufacturing highlights an application of ML techniques to enhance detection accuracy, optimize manufacturing processes, and support efficient decision-making in quality control and defect prevention.

In this study, Tsai and Lee (2020) suggested using depth-wise separable convolutions as the basis for a reduced-weight architecture classifier. The real-world Wafer Map dataset (WM-811K) is used to validate the whole study. In the test set, the accuracy is 96.63%. The test procedure is the primary source of manufacturing cost in the IC design process. To ascertain the process's current state, existing tests depend on the engineer doing extra analysis on the testing result data. Consequently, it may need more time and cannot quickly implement process improvements. An essential component of semiconductors is wafer map defect recognition. Engineers can swiftly determine the kind of failure because to the wealth of information included in wafer maps[13].

In this study, Tziolas et al. (2022) the suggested CNN-based model is used on the publicly available but very unbalanced industrial dataset WM-811K and makes use of a number of pre-and post-processing techniques. To address imbalance, a solution is suggested that treats each class independently by using distinct processing methods for data augmentation, splitting, and down-sampling depending on the sample size. Outperforming competing models in the relevant literature, the suggested model identifies the vast majority of classes with a 95.3% accuracy and a 93.78% macro F1-score [14].

In this study, Zhai, Shi and Zeng (2023) puts forward a domain-specific Auto ML (called an Auto Classifier) for anomaly detection and defect identification to enable end-to-end machine learning by self-learning the best models for semiconductor flaws. The first step is feature engineering, which entails cleaning the data, extracting features using interpolation techniques, and then selecting the best features to use from the SECOM dataset. the focused loss is included in cutting-edge classifiers XGBoost and LightGBM, which aim to resolve the pervasive data imbalance in the manufacture of semiconductor wafers. After going from an overall F1 score of 85% to 92.3%, the comparison and algorithm selection show that focus loss is beneficial[15].

In this study, Park et al. (2024) analyzed and preprocessed data to determine the likelihood of excellent and faulty products in semiconductor production, with the goal of increasing yield while decreasing costs via the use of ML. This is accomplished by using the SECOM dataset and carrying out preparation procedures such managing missing values, reducing dimensionality, resampling to fix class imbalances, and scaling. Lastly, measures like the geometric mean (GM) were used to compare and contrast six ML models that were trained on unbalanced data using different combinations of preprocessing strategies. With the goal of decreasing training and prediction timeframes, this study differs from others in that it suggests ways to decrease the amount of features used in ML. This work goes a step further by segregating the training and test datasets prior to analysis and preprocessing, which eliminates data leaking during this stage. Using oversampling approaches (not including KM SMOTE) leads to more equitable class categorization, according to the findings. The most effective combination was SVM with ADASYN and MaxAbs scaling, which achieved an accuracy of 85.14% and a GM of 72.95%, respectively[16].

In this study, Tsai and Lee (2020) showcase a process for improving wafer map data and categorizing defects. The CNN encoder-decoder is the basis of the data augmentation, while depth-wise separable convolutions provide the basis of the classification. The first dataset is the open-source WM-811K dataset, while the second is a custom-built one that was jointly created with a company in Taiwan. Two models are trained using mobilenetV1 and V2 for the two datasets. Testing houses with high production volumes might benefit greatly from the lightweight deep convolution's ability to decrease model parameters and computations. Their suggested approach may decrease the amount of computation by 75% and 95% and the number of parameters by 30% and 95% on two separate data sets, respectively. There is a 93.95% test accuracy in the first dataset. There is an accuracy of 87.04% in the second dataset. Achieving 97.01% and 95.09% accuracy, respectively, after data augmentation [17].

In this study, Bhatnagar, Arora and Chaujar (2022) showcase and contrast several transfer learning techniques that can label the wafer map with the appropriate faults. Their trials evaluate the efficiency of several models based on transfer learning, such as VGG19, MobileNet, ResNet, and DenseNet, using the WM811K dataset, which represents a real-world wafer map. Eight different types of wafer map defects are included in the dataset. They have grouped all of the flaws into four different types, which will be covered later on in the study. The results show that VGG19 attained the best accuracy of 95.56% on the test data, according to their trials[18].

In this study, Chen et al. (2023) suggested functionality. Higher fine-grained and richer features may be obtained via a fusion module to preserve crucial information and capture significant texture characteristics. The results of the last trials show that MFFP-Net has great generalizability and can achieve state-of-the-art performance on the real-world dataset WM-811K. This provides a realistic way for the chip manufacturing industry to boost yield rate[19].

Table I presents the identified research gaps in previous studies on fault detection in semiconductor manufacturing using machine learning techniques, highlighting key areas that require further exploration and advancement.

© JGREC 2025, All Rights Reserved

TABLE I. SUMMARY OF THE RELATED WORK BASED ON FAULT DETECTION IN SEMICONDUCTOR MANUFACTURING USING MACHINE LEARNING APPROACHES.

References	Methodology	Dataset	Performance	Limitations & Future Work
[13]	CNN-based classifier with reduced-weight architecture using depthwise separable convolutions	WM-811K	96.63% accuracy	Limited focus on handling data imbalance; future work can explore adaptive weighting techniques
[14]	Pre- and post-processing techniques with CNN-based model, handling imbalance through down-sampling, splitting, and augmentation	WM-811K	95.3% accuracy, 93.78% macro F1-score	Requires further investigation into advanced data augmentation techniques and handling rare defect classes
[15]	Auto ML-based system (Auto Classifier) integrating feature engineering and focal loss with LightGBM and XGBoost.	SECOM	F1-score improved from 85% to 92.3%	Needs further validation on other industrial datasets and exploration of deep learning-based approaches
[16]	Machine learning models with preprocessing steps (missing value handling, dimensionality reduction, resampling, scaling)	SECOM	SVM with ADASYN & MaxAbs scaling achieved 85.14% accuracy, 72.95% GM	Further optimization of preprocessing and feature selection needed; alternative oversampling methods could be explored.
[17]	CNN encoder-decoder for data augmentation, Depth wise separable convolutions for classification	WM-811K (open dataset) & Taiwan company dataset	Accuracy: 93.95% → 97.01% (WM-811K) 87.04% → 95.09% (Taiwan dataset) Reduction in parameters: 30%-95% Reduction in computation: 25%-75%	Potential limitations in generalizability, real-time industrial deployment challenges, impact of augmentation on unseen defects
[18]	Transfer learning-based classification using VGG19, MobileNet, ResNet, DenseNet	WM-811K	VGG19 achieved highest accuracy of 95.56%	Merging defect classes may affect generalizability; future work can focus on fine-grained defect classification.
[19]	Feature fusion module (MFFP-Net) to enhance fine-grained feature extraction	WM-811K	96.71% accuracy	Generalization to other semiconductor datasets needs validation; further improvements in interpretability could be explored.

III. METHODOLOGY

This study aims to improve fault detection accuracy in semiconductor manufacturing using machine learning techniques. By applying a CNN with data augmentation to the WM811K dataset, the research focuses on effectively identifying wafer defects and addressing challenges like data imbalance and overfitting. Performance is evaluated through metrics like F1-score, recall, accuracy, and precision, with comparisons to models like VGG19 and XGBoost. The evaluation's findings will contribute to the improvement of the

methodology for improving fault detection in semiconductor manufacturing involves utilizing the WM811K dataset, which contains 811,457 wafer images, including 172,950 manually labeled images across eight defect categories. Preprocessing steps include handling missing values, noisy data, and inconsistencies, followed by data augmentation using an autoencoder for dimensionality reduction and noise introduction to expand the dataset and enhance model performance. The dataset is split into 80% for training and 20% for testing. The CNN architecture uses convolutional layers, activation functions (ReLU), fully connected layers, and pooling layers to classify and extract features. The model's performance is evaluated using confusion matrix analysis and key measures like as F1-score, recall, precision, and accuracy. Comparative analysis with VGG19 and XGBoost demonstrates the superior fault detection accuracy of the CNN-AUG model, which effectively captures intricate spatial features.

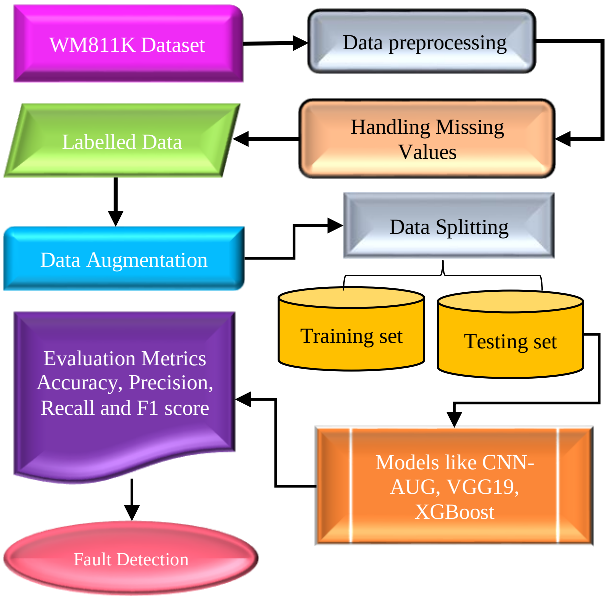

Fig. 1. Flowchart depicting the machine-learning based fault detection process in semiconductor manufacturing.

The following steps of a flowchart based on fault detection in semiconductor manufacturing briefly explained in below:

A. Data Collection and preprocessing

Utilized the WM811K semiconductor dataset for their experiments. There are 811,457 photos of wafers in this extensive collection, which also includes information like batch numbers, indices, and core dimensions. The dataset was collected from 47,543 physical lots from a FAB, with each lot consisting of 25 wafers. However, although 47,543 lots would yield 1,557,325 wafers, the dataset contains only 811,457 wafer images. The process of transforming unstructured data

© JGREC 2025, All Rights Reserved

into a more usable format is known as data preparation. Before running the method on the dataset, it undergoes preprocessing to remove any inconsistent or missing values, noise, or other issues. The manually labeled portion of the dataset comprises 172,950 images with 8 distinct labels (0-7). Due to sensor malfunctions or other unidentified problems, not every batch has 25 wafer images. As a result, 172,950 wafers were labeled, while the remaining 78.7% of the wafers were unlabeled. Among the labeled wafers, 25,519 wafers (3.1%) were defective, while the remaining 147,431 wafers were intact. This highlights the scarcity of defective samples available for training the model. The distribution of defects is as follows: Center: 4294, Donut: 555, Edge-Loc: 5189, Edge-Ring: 9680, Loc: 3593, Random: 866, Scratch: 1193, Near-full: 149. The dataset contains 18.2% normal, defect-free wafers (labeled as 8), whereas the other labels (0–7) indicate different patterns of defects. Data augmentation is used to improve size of the dataset. Data augmentation is implemented through the use of an autoencoder model for dimensionality reduction into the latent space, where noise is introduced, and the wafer map data is reconstructed, ultimately increasing the sample size and improving the model's performance.

B. *Data Splitting*

The dataset is partitioned into two subsets to enhance fault detection accuracy: 80% is allocated for model training, facilitating the development of an effective fault detection system, while the remaining 20% is reserved for testing to assess the model's performance and generalizability in semiconductor manufacturing.

C. *Implementation of Classification Using CNN with Data Augmentation*

There are a wide variety of CNN topologies. Certain layers are required for integrating input before classification, such as pooling layers, convolution layers with activation functions, and a completely connected layer or layers. Next, the feature maps were retrieved by the convolution layers. To make the subsequent layers' processing simpler, the pooling layers concentrated the feature maps using techniques like maximizing or average values inside a certain frame [20]. The categorization process begins after these layers and continues with a completely connected layer or layers. They will not restate the arguments presented in the literature that CNN is successful in feature extraction. Each of the tagged images is used as input during the training phase when the CNN is expected to deliver the correct label. Depending on how many epochs or the convergence threshold is chosen, the complete training process may take a lengthy time.

Weight matrices w, and bias matrices b are the building blocks of each convolution layer; the training process updates and initializes these matrices. Equation 1 would display the results of every convolution layer:

$$x_i^l = \sigma\left(\sum_{i\epsilon FM_j} w_i x_j^{l-1} \times k_{ij}^l + b_j^l\right) \tag{1}$$

where σ(.) denotes an activation function for every link from i to j, l denotes a layer, FM denotes a feature map, k denotes a convolution kernel, b denotes a bias, and. The activation function is, generally, a ReLU function, where ReLU(x) = max (0, x), and

$$ReLU(x) = \max(0, x)\,, and$$
$$\frac{d}{dx}ReLU(x) = \begin{Bmatrix} 1, if\ x > 0 \\ 0, otherwise \end{Bmatrix} \tag{2}$$

D. *Key Performance Evaluation Metrics*

The efficacy of the deployed model may be evaluated with the help of the performance measures. Several metrics provide light on how well ML models are doing in general, including accuracy, recall, F1-score, and precision.

1) *Confusion matrix*

Statistical classification is the main use for the confusion matrix, often known as the error matrix. A particular table arrangement enables the visualization of an algorithm's performance. One need not assume a relationship between the two values if the projected value appears in the row of the matrix and the actual value in the column. False positive (FP), true negative (TN), true positive (TP), and true negative (FN) are the four cells that make up the output matrix. TP indicates that the actual and anticipated values are positively related to one another; TN indicates a favorable correlation between the observed and anticipated values; A negative correlation among the actual and anticipated values of the model is represented by FP; and the negative correlation among the actual and projected values is represented by FN. The following performance matrix discussed below in detail:

2) *Accuracy*

A classifier's accuracy is defined as the frequency with which it makes right predictions from the whole dataset, as shown in Equation (3).

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \tag{3}$$

3) *Precision*

Precision is defined as the ratio of total TP to the total TP plus FP. The formula is presented in Equation (4).

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

4) *Recall*

The amount of FN included in a prediction mixture is the main focus of recall [21]. Equation (5) calculates the recall, which is often called the sensitivity or true positive rate:

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

5) *F1-score*

Equation (6) defines the F1-score, which is the harmonic mean of recall and precision.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

These performance metrics are utilized to assess an efficacy of a model by examining its outcomes on the test set.

IV. Result Analysis And Discussion

The CNN-AUG, VGG19, and XGBoost models proved to be the most successful in detecting errors in semiconductor production when tested against other machine learning algorithms. In terms of detecting complex spatial characteristics, the CNN-AUG model performs better than standalone models, according to the comparison. A variety of visual representations, including the confusion matrix, training/validation accuracy/loss plots, and performance measures including recall, accuracy, precision, and F1-score, provide light on how well the models improved fault detection accuracy.

TABLE II. CNN-AUG Model Performance based on fault detection using WM811K Dataset.

Metrics	CNN-AUG
Accuracy	98.56
Precision	98.77
Recall	98,78
F1 Score	98.77

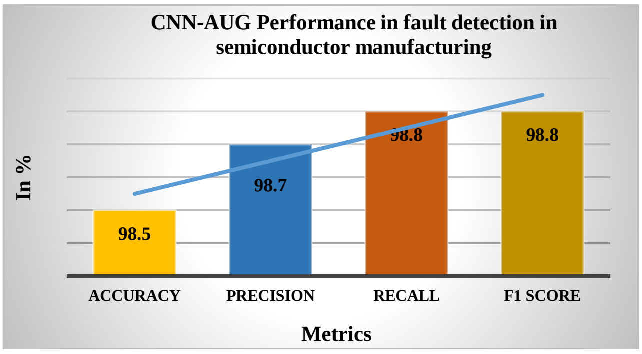

Fig. 2. CNN-AUG Performance in fault detection in semiconductor manufacturing.

Table II and Figure 2 displays the results of CNN-AUG's fault detection performance in semiconductor production, as measured by F1 score, recall, accuracy, and precision. The outcomes indicate a high level of performance across all metrics, with F1 score demonstrating the highest value at 98.8%. This suggests that CNN-AUG is effective in identifying and classifying faults in semiconductor manufacturing processes.

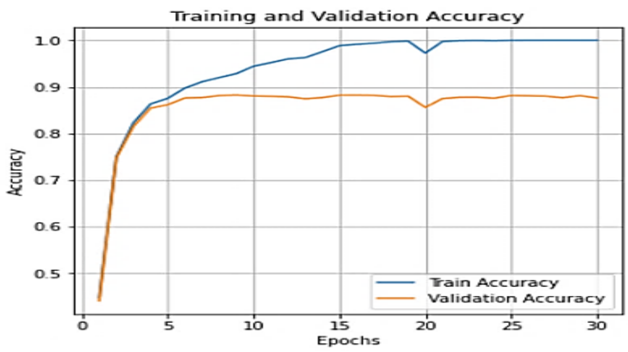

Fig. 3. Training and validation loss of CNN-AUG model

Figure 3 displays the accuracy of an ML model that was trained and validated to identify faults in semiconductor production. Validation accuracy reaches a ceiling when model performance on training data rises continuously. This raises the possibility of overfitting, a phenomenon in which a model becomes too reliant on its training data and fails to adequately adapt to more complex manufacturing problems. Fixing this overfitting is the first step toward making semiconductor manufacturing defect detection systems more accurate and reliable.

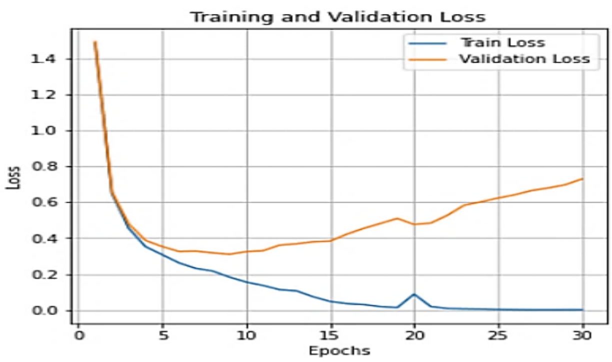

Fig. 4. Training and validation loss of CNN-AUG model

Figure 4 illustrates the training and validation loss of a machine learning model used for fault detection in semiconductor manufacturing. The model is learning as the training loss goes down, which is a good sign. The validity loss, however, drops at first before beginning to rise above a certain threshold. Overfitting occurs when a model learns patterns in its training data that do not translate well to novel, unseen data; this might be happening if the model is not careful.

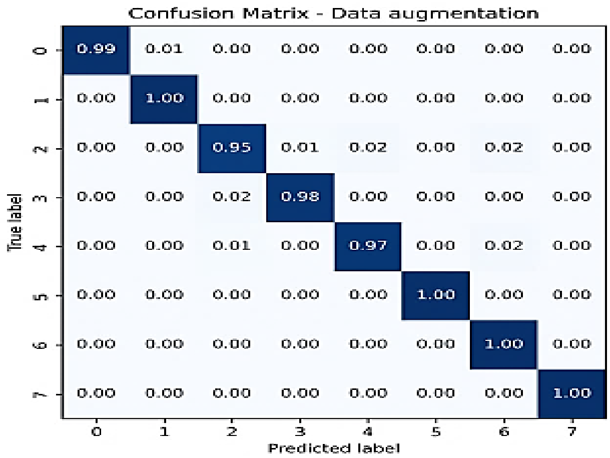

Fig. 5. Confusion matrix for fault detection in semiconductor manufacturing

Figure 5 visualizes the confusion matrix performance of a fault detection model in semiconductor manufacturing, likely using data augmentation. The diagonal values, close to 1, indicate high accuracy in classifying faults into their respective categories. Off-diagonal values close to 0 suggest minimal misclassification. The matrix demonstrates the model's effectiveness in correctly identifying different fault types, which is crucial for improving yield and reducing defects in semiconductor production.

TABLE III. COMPARISON EVALUATION OF ML MODELS PERFORMANCE FOR FAULT DETECTION

Metrics	Accuracy	Precision	Recall	F1 Score
CNN-AUG	98.56	98.77	98.78	98.77
VGG19 [22]	89.3	90.15	89.46	89.49
XGBoost [23]	89	87	89	88

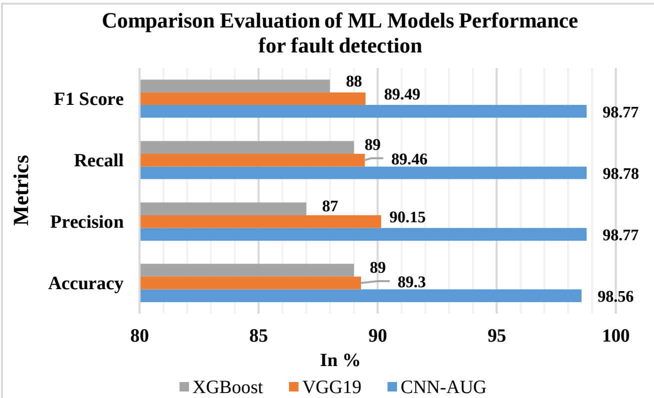

Fig. 6. Comparative Evaluation of ML Models for Load Forecasting

The Table III and bar graph in the Figure 6 compares four key metrics: F1 Score, Recall, Precision, and Accuracy, across three models: XGBoost, VGG19, and CNN-AUG. The scores

© JGREC 2025, All Rights Reserved

are represented as percentages, ranging from 80% to 100%. CNN-AUG consistently outperforms the other two models across all metrics, achieving nearly perfect scores. VGG19 shows moderate performance, while XGBoost has the lowest scores. The visual representation facilitates easy comparison and highlights the superior efficacy of the CNN-AUG model in fault detection.

V. Conclusion And Future Scope

Improving fault detection precision in semiconductor production is essential for enhancing product quality and minimizing faults. Utilizing machine learning methodologies, especially deep learning, can lead to substantial enhancements in defect classification. Of the assessed models, CNN-AUG exhibited superior performance, attaining 98.56% Accuracy, 98.77% Precision, 98.78% Recall, and 98.77% F1-score, thereby validating its efficacy in identifying semiconductor flaws. The implementation of preprocessing strategies, such as addressing missing values, utilizing labeled data, and applying data augmentation, enhanced model resilience. These findings highlight the capability of AI-driven solutions in enhancing semiconductor production processes. However, the model's reliance on extensive labeled datasets and data augmentation techniques may limit its deployment in environments with scarce resources. Additionally, further investigations are necessary to validate the model's scalability and adaptability across various semiconductor manufacturing conditions.

Future research may investigate hybrid deep learning models that combine CNNs with transformers to improve feature extraction. Employing semi-supervised or unsupervised learning can mitigate issues associated with insufficient labelled data. Furthermore, real-time fault detection with edge computing and model interpretability methods such as SHAP can improve industry adoption. Subsequent optimizations will enhance the creation of more precise, scalable, and intelligent defect detection systems, hence augmenting semiconductor manufacturing efficiency.

References

[1] H. Sarpana Chandu, "Robust Control of Electrical Machines in Renewable Energy Systems: Challenges and Solutions," *Int. J. Innov. Sci. Res. Technol.*, vol. 09, no. 10, pp. 594–602, Oct. 2024, doi: 10.38124/ijisrt/IJISRT24OCT654.

[2] P. Tchatchoua, G. Graton, M. Ouladsine, and J. F. Christaud, "Application of 1D ResNet for Multivariate Fault Detection on Semiconductor Manufacturing Equipment †," *Sensors*, 2023, doi: 10.3390/s23229099.

[3] H. S. Chandu, "Enhancing Manufacturing Efficiency: Predictive Maintenance Models Utilizing IoT Sensor Data," *IJSART*, vol. 10, no. 9, 2024.

[4] V. Panchal, "Mobile SoC Power Optimization: Redefining Performance with Machine Learning Techniques," *Int. J. Innov. Res. Sci. Eng. Technol.*, vol. 13, no. 12, Dec. 2024, doi: 10.15680/IJIRSET.2024.1312117.

[5] H. S. Chandu, "A Survey of Semiconductor Wafer Fabrication Technologies: Advances and Future Trends," *Int. J. Res. Anal. Rev.*, vol. 10, no. 04, pp. 344–349, 2023.

[6] P. K. Vishwakarma and M. R. Suyambu, "An Analysis of Engineering , Procurement And Construction (EPC) -Contracts Based on Renewable Energy," *IJSART*, vol. 10, no. 10, pp. 26–35, 2024.

[7] M. Saqlain, Q. Abbas, and J. Lee, "A Deep Convolutional Neural Network for Wafer Defect Identification on an Imbalanced Dataset in Semiconductor Manufacturing Processes," *IEEE Trans. Semicond. Manuf.*, vol. PP, p. 1, 2020, doi: 10.1109/TSM.2020.2994357.

[8] A. P. A. S. and N. Gameti, "Digital Twins in Manufacturing: A Survey of Current Practices and Future Trends," *Int. J. Sci. Res. Arch.*, vol. 13, no. 1, pp. 1240–1250, 2024.

[9] V. Panchal, "Energy-Efficient Core Design for Mobile Processors : Balancing Power and Performance," pp. 191–201, 2024.

[10] Muthuvel Raj Suyambu and Pawan Kumar Vishwakarma, "State-of-Art Techniques for Photovoltaic (PV) Power Systems and their Impacts," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 4, no. 3, pp. 381–389, Oct. 2024, doi: 10.48175/IJARSCT-19956.

[11] Z. M. Çınar, A. Abdussalam Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, "Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0," *Sustainability*, vol. 12, no. 19, 2020, doi: 10.3390/su12198211.

[12] A. J. Rahul Dattangire, Ruchika Vaidya, Divya Biradar, "Exploring the Tangible Impact of Artificial Intelligence and Machine Learning: Bridging the Gap between Hype and Reality," *2024 1st Int. Conf. Adv. Comput. Emerg. Technol.*, pp. 1–6, 2024.

[13] T.-H. Tsai and Y.-C. Lee, "Wafer Map Defect Classification with Depthwise Separable Convolutions," in *2020 IEEE International Conference on Consumer Electronics (ICCE)*, 2020, pp. 1–3. doi: 10.1109/ICCE46568.2020.9043041.

[14] T. Tziolas *et al.*, "Wafer Map Defect Pattern Recognition using Imbalanced Datasets," in *13th International Conference on Information, Intelligence, Systems and Applications, IISA 2022*, 2022. doi: 10.1109/IISA56318.2022.9904402.

[15] W. Zhai, X. Shi, and Z. Zeng, "Adaptive Modelling for Anomaly Detection and Defect Diagnosis in Semiconductor Smart Manufacturing: A Domain-specific AutoML," in *2023 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, 2023, pp. 198–203. doi: 10.1109/CIS-RAM55796.2023.10370014.

[16] H.-J. Park, Y.-S. Koo, H.-Y. Yang, Y.-S. Han, and C.-S. Nam, "Study on Data Preprocessing for Machine Learning Based on Semiconductor Manufacturing Processes," *Sensors*, vol. 24, no. 17, p. 5461, Aug. 2024, doi: 10.3390/s24175461.

[17] T.-H. Tsai and Y.-C. Lee, "A Light-Weight Neural Network for Wafer Map Classification Based on Data Augmentation," *IEEE Trans. Semicond. Manuf.*, vol. 33, no. 4, pp. 663–672, 2020, doi: 10.1109/TSM.2020.3013004.

[18] P. Bhatnagar, T. Arora, and R. Chaujar, "Semiconductor Wafer Map Defect Classification Using Transfer Learning," in *2022 IEEE Delhi Section Conference (DELCON)*, 2022, pp. 1–4. doi: 10.1109/DELCON54057.2022.9753436.

[19] Y. Chen, M. Zhao, Z. Xu, K. Li, and J. Ji, "Wafer defect recognition method based on multi-scale feature fusion," *Front. Neurosci.*, vol. 17, p. 1202985, 2023.

[20] J.-C. Chien, M.-T. Wu, and J.-D. Lee, "Inspection and classification of semiconductor wafer surface defects using CNN deep learning networks," *Appl. Sci.*, vol. 10, no. 15, p. 5340, 2020.

[21] K. H. Ang *et al.*, "Optimized Convolutional Neural Network Towards Effective Wafer Defects Classification," 2024.

[22] H. Zheng, S. W. A. Sherazi, S. H. Son, and J. Y. Lee, "A Deep Convolutional Neural Network-Based Multi-Class Image Classification for Automatic Wafer Map Failure Recognition in Semiconductor Manufacturing," *Appl. Sci.*, vol. 11, no. 20, 2021, doi: 10.3390/app11209769.

[23] A. Presciuttini, A. Cantini, and A. Portioli-Staudacher, "Advancing Manufacturing with Interpretable Machine Learning: LIME-Driven Insights from the SECOM Dataset," in *IFIP International Conference on Advances in Production Management Systems*, 2024, pp. 286–300.

© JGREC 2025, All Rights Reserved